ZooKeys 746: 105–121 (2018)
doi: 10.3897/zookeys.746.13502
http://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Re-evaluation of the discriminatory power of DNA barcoding on some specimens of African Cyprinidae (subfamilies Cyprininae and Danioninae)

Mariam I. Adeoba[1], Ronny Kabongo[2],
Herman Van der Bank[1], Kowiyou Yessoufou[3]

1 *Department of Zoology, University of Johannesburg, Kingsway Campus PO Box 524, Auckland Park 2006, South Africa* **2** *African Centre for DNA Barcoding, University of Johannesburg, Kingsway Campus, PO Box 524, Auckland Park 2006, South Africa* **3** *Department of Geography, Environmental management and Energy studies, University of Johannesburg, Kingsway Campus PO Box 524, Auckland Park 2006, South Africa*

Corresponding author: *Mariam I. Adeoba* (mariamsalami@yahoo.co.uk)

Academic editor: *N. Bogutskaya* | Received 2 May 2017 | Accepted 14 December 2017 | Published 26 March 2018

http://zoobank.org/6627B542-F87A-423F-B5CD-C04692D16ED9

Citation: Adeoba MI, Kabongo R, Van der Bank H, Yessoufou K (2018) Re-evaluation of the discriminatory power of DNA barcoding on some specimens of African Cyprinidae (subfamilies Cyprininae and Danioninae). ZooKeys 746: 105–121. https://doi.org/10.3897/zookeys.746.13502

Abstract

Specimen identification in the absence of diagnostic morphological characters (e.g., larvae) can be problematic even for experts. The goal of the present study was to assess the performance of COI in discriminating specimens of the fish family Cyprinidae in Africa, and to explore whether COI-phylogeny can be reliably used for phylogenetic comparative analysis. The main objective was to analyse a matrix of COI sequences for 315 specimens from 15 genera of African Cyprinidae using various distance-based identification methods alongside multiple tests of DNA barcode efficacy (barcode gap, species monophyly on NJ tree). Some morphological and biological characters were also mapped on a COI-phylogeny reconstructed using Maximum Parsimony. First, the results indicated the existence of barcode gaps, a discriminatory power of COI ranging from 79 % to 92 %, and that most nodes form well-supported monophyletic clades on an NJ tree. Second, it was found that some morphological and biological characters are clustered on the COI-phylogeny, and this indicates the reliability of these characters for taxonomic discrimination within the family. Put together, our results provide not only an additional support for the COI as a good barcode marker for the African Cyprinidae but it also indicate the utility of COI-based phylogenies for a wide spectrum of ecological questions related to African Cyprinidae.

Copyright *Mariam I. Adeoba et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords
BRONX algorithm, character variations, COI, specimen identification

Introduction

Cyprinidae is the most diverse family of freshwater fishes (Nelson et al. 2006; Imoto et al. 2013) with 377 genera and over 3000 described species (Eschmeyer and Fong 2015; Froese and Pauly 2017). Species of this family are mostly found in Africa, Europe, Asia, and North America (Thai et al. 2007). In Africa, recent studies of the family have identified 24 genera and 539 species (Yang et al. 2015; Vreven et al. 2016). Some species of the family are of economic importance in aquaculture, angling, fisheries, aquarium trade and many serve as an essential source of protein for humans in addition to their high values in recreational fisheries (Skelton 2001; Thai et al. 2007; Collins et al. 2012).

Traditionally, external morphological and osteological characteristics have been used to differentiate species within the subfamilies Cyprininae and Danioninae (Zhou 1989; Chen et al. 2009; Yang et al. 2010; Liao et al. 2011; Nelson et al. 2016). For example, diagnostic characters such as a spinous anal-fin ray in some Cyprininae (Yang et al. 2010), interhyal not ossified (Liao et al. 2011) and an extended anal fin in mature males of some Danioninae (Stiassny et al. 2006) are used for specimen identification in both subfamilies. Additionally, there are key morphological features that distinguish the males from the females, including a brighter breeding colour, longer fins and presence of the tubercles on the body and head in some African genera (Skelton 2001). Similarly, morphological features such as the presence/absence of barbels, the number of barbels, as well as barbel type, pattern of innervation, and barbel position have been used to differentiate species within and between genera of the subfamilies Cyprininae sensu lato and Danioninae sensu lato (Howes 1991; Skelton 2001).

The taxonomy of the family has been a topic debated in several studies (e.g., Howes 1991; Cavender and Coburn 1992; Briolay et al. 1998; Zardoya and Doadrio 1999; Gilles et al. 2001; Yang et al. 2015; Ren and Mayden 2016). Some studies have explored the phylogeny of this family at subfamily and genus levels using both mitochondrial and nuclear genes (Zardoya and Doadrio 1999; Simons et al. 2003; Stiassny and Getahun 2007; Tang et al. 2010; Yang et al. 2010; Zheng et al. 2012; Yang et al. 2015). Specifically in Africa, most cyprinid species were previously assigned to the subfamily Cyprininae (Tsigenopoulos et al. 2002). The former genus *Barbus* forms a large polyphyletic group of more than 800 species across the world and 300 species across Africa (Skelton et al. 1991). Early studies used, in addition to morphological characteristics, the ploidy level to reorganise the genus *Barbus* sensu lato in Africa (Agnèse et al. 1990; Oellermann and Skelton 1990; Güegan et al. 1995; Berrebi et al. 1996; Machordom and Doadrio 2001; Tsigenopoulos et al. 2002). As a result, some African *Barbus* from northern and southern Africa have been regrouped into genera such as *Luciobarbus*

and *Pseudobarbus* (Swartz et al. 2009; Tsigenopoulos et al. 2010) and other species now belong to genus *Labeobarbus* (Oellermann and Skelton 1990; Berrebi et al. 1996; Machordom and Doadrio 2001).

Similarly, the recent molecular and morphological work of Yang et al. (2015) on subfamily Cyprininae had led to a major reclassification and name changes in the global Cyprinidae. This reclassification has since been adopted in some recent works (Armbruster et al. 2016; Decru et al. 2016; Skelton 2016; Vreven et al. 2016). As a result, some genera within the African Cyprinidae are now subfamilies (e.g., Cyprininae, Danioninae and Leuscininae) with few species belonging to non-specified subfamilies (Suppl. material 1). Presently, the African Cyprininae is grouped into four tribes including Barbini, Smiliogastrini, Torini and Labeonini (Yang et al. 2015). The tribe Barbini includes genera such as *Luciobarbus*, *Barbopsis*, *Caecobarbus* and *Coptostomabarbus* and the Smiliogastrini includes the genera *Barbodes*, *Barboides*, *Clypeobarbus*, *Enteromius* and *Pseudobarbus*. The former African diploid '*Barbus*' is now reclassified within the genus *Enteromius* (Yang et al. 2015; Armbruster et al. 2016) and the South African tetraploid *Barbus* has been elevated to genus '*Pseudobarbus*' (Yang et al. 2015; Skelton 2016), although Schmidt and Bart (2015) suggested a revision for genus *Pseudobarbus* to clarify those with inverted comma. Additionally, the former African *Varicorhinus* was reassigned to *Labeobarbus* in the tribe Torini (Beshera et al. 2016; Skelton 2016; Vreven et al. 2016). Yang et al. (2015) also suggested *Sanagia velifera* Holly, 1926 to be grouped with the genus *Labeobarbus*. The tribe Labeonini includes the genera *Labeo*, *Garra* and *Prolabeo* (Rainboth et al. 2012; Yang et al. 2012, 2015). In such context of taxonomic debate around the family Cyprinidae (Yang et al. 2015), it becomes necessary to question whether the ongoing global campaign of DNA barcoding can play a role at least in assigning specimen to their corresponding taxa. The DNA barcoding approach has been employed to complement or refine morphological species identification (Kochzius et al. 2010; Pereira et al. 2011; Chen et al. 2015). DNA barcoding is based on the use of a short standardised cytochrome c oxidase subunit I (COI) sequence to distinguish between animal species (Hebert et al. 2003; 2004). It has gained worldwide support because it is rapid, cost-effective (but see Stein et al. 2014), and applicable to species identification across the animal kingdom (e.g., Hebert et al. 2003; Ward et al. 2005; Van der Bank et al. 2013; Sethusa et al. 2014; Decru et al. 2016; Nigro et al. 2016). In particular, Decru et al. (2016) clearly demonstrated, using DNA barcoding, how knowledge of the African fish species diversity can be greatly improved, but they focused only on the Congo Basin region in Central Africa.

The present study uses a broader sampling of the African Cyprinidae and integrates morphology and ploidy data to further assess the effectiveness of DNA barcoding in discriminating specimens within the family. Specifically, the aim was to: (i) test the reliability of COI as a DNA barcode for the African Cyprinidae based on barcode gap, various distance methods, and the Rosenberg test of species monophyly; and (ii) map six traits including five morphological characters and ploidy level onto a COI-based phylogeny of the African Cyprinidae.

Materials and methods

Sample collections

First, 584 COI sequences of the African Cyprinidae specimens were retrieved from the Barcode of Life Database (BOLD; www.boldsystems.org) and GenBank/EBI (www.ncbi.nlm.nih.gov/nuccore). Some of the sequences from BOLD had been generated from our group (African Centre for DNA Barcoding) (Suppl. material 1). Second, for the purpose of the present study, an additional set of 55 new sequences of southern African specimens were generated to create a total DNA matrix of 639 specimens consisting of 15 of the 24 genera of African Cyprinidae. Sequences of the 55 specimens are made available on BOLD and GenBank/EBI. The BOLD identification numbers, voucher information, GenBank accession numbers, and species authorities for all species analysed in this study are presented in Suppl. material 2. Localities, images and additional information are also available on BOLD. It should also be noted that, as a result of the ongoing taxonomic revision and debates around this family, some of the African species names have been altered in FishBase but are yet to be updated on BOLD and GenBank. Therefore, for this study the old names were retained in our analysis (see Suppl. material 2; but new names are adopted in Figures 4 and 5). All the species analysed in the present study are those that have accession numbers in Suppl. material 1.

DNA extraction, amplification, and sequencing of COI

The 55 new COI sequences mentioned above were generated following the manufacturers' recommended protocol developed from NucleoSpin® Tissue kit (Macherey-Nagel). The sequence amplification (PCR) was done in accord with Hajibabaei et al. (2005). Specifically, PCR reactions were done in a total volume of 25 µL. The master mix consisted of 12.5 µL of top taq, 0.8 µL of BSA, 0.3 µL of both primers and 10.1 µL of dH2O. The DNA templates prepared for the PCR amplification ranged from 1–3 µL, depending on the strength and quality of DNA products visualized from the agarose gel. The PCR conditions were as follows: initial melting for 2 mins at 95 oC, denaturation at 94 oC for 0.5 min, annealing at 52 oC for 0.5 min, extension at 72 oC for 1 min followed by a final extension at 72 oC for 10 mins (35 cycles) and a hold at 4 oC (Steinke and Hanner 2011). The primer pair used was COI-Fish. F1 5'-TCAACCAACCACAAAGACATTGGCAC-3' and COI-Fish.R1 3'-TAGAC TCTG GGTGGCCAAAGAATCA-5'.

After the amplification, PCR products were visualised on 1.5% agarose gels. Visible products were cleaned using silica column kits, viewed again on agarose gels, and selected for cycle sequencing. Sequencing of COI region was done following the standard protocols of the Canadian Centre for DNA Barcoding (CCDB). Sequences were aligned using Multiple Sequences Comparison by Log-Expectation (MUSCLE vs. 3.8.31; Edgar 2004) and exported as a NEXUS file.

DNA barcoding analysis

Because some DNA sequences available on public repositories are not reliable (Nilsson et al. 2006), we first used the BRONX algorithm (Barcode Recognition Obtained with Nucleotide eXpose´s; Little 2011) to reanalyse all sequences retrieved from BOLD and GenBank/EBI to refine the dataset prior to our DNA barcoding analysis. Based on the BRONX analysis, we removed from our dataset (of 639 sequences) sequences that are questionable, for a number of reasons, including shared haplotypes between species, shorter sequences, and incomplete identification, etc. Also, species with no duplicates (singletons) were excluded, and as a result, the total samples included in our DNA barcoding analysis comprise 315 sequences for 86 species representing 14 out of the 24 (58 %) recognised genera in Africa (Suppl. material 2).

All barcoding analysis was conducted in the R package SPIDER (species identity and evolution in R) vs. 1.1-1 (Brown et al. 2012) following three criteria: barcoding gap, discriminatory power, and tree based analysis for species monophyly. Two techniques were used in evaluating the "DNA barcode gap" (Meyer and Paulay 2005). Firstly, the mean, median, and range of intraspecific genetic distances were compared against interspecific genetic distance (Meier et al. 2008). Secondly, the approach of Meier et al. (2006) was used to assess barcode gap. This involves matching the lowest interspecific distance against the highest intraspecific distance. Genetic distances were calculated using the Kimura 2-parameter (K2P) model (Kimura 1980).

The discriminatory power of the COI gene was tested with three methods: Best Close Match, Near Neighbour and the BOLD identification (threshID) (Meier et al. 2006; Collins et al. 2012). A good barcode should exhibit a high rate of correct species identification. Prior to the analysis, the optimised threshold for specimen identification was first determined using the R function *localMinima* (Brown et al. 2012) and then applied in the Best Close Match and Near Neighbour identification. The identification success of the traditional 1% threshold of BOLD was additionally tested in comparison to bestCloseMatch (Brown et al. 2012).

To test for species monophyly, a tree based analysis using Rosenberg's (2007) probability of reciprocal monophyly and a Neighbour-Joining (NJ) phylogram was constructed (Rosenberg 2007). For this purpose, our default was set to be false for singletons and our tree rooted on the longest branch with labels corresponding to species vector (Brown et al. 2012).

Phylogenetic reconstruction and character mapping

A DNA matrix of 315 COI aligned sequences and three outgroups (Suppl. material 3) was formed, and this matrix used to assemble a phylogeny based on Maximum Parsimony (MP) using PAUP* v4.0b 10 (Swofford 2002) with heuristic searches and 1,000 random-addition sequence replicates and tree-bisection-reconnection branch swapping. The following outgroups were chosen from similar past studies: *Moxostoma breviceps* (Cope,

Table 1. Tests of barcoding identification accuracy with numbers (N) and percentages (%) of each score Near neighbour, Best Close Match and Bold criteria.

Methods	Near neighbour		Best Close match				Bold criteria			
Scores	False	True	Ambiguous	Correct	Incorrect	No ID	Ambiguous	Correct	Incorrect	No ID
N	25	290	10	278	17	10	57	250	1	7
(%)	7.9	92.1	3.2	88.2	5.4	3.2	18.1	79.4	0.3	2.2

1870) (De Graaf et al. 2007), *Pseudorasbora parva* (Temminck & Schlegel, 1846) and *Gyrinocheilus aymonieri* (Tirant, 1883) (He et al. 2008) (Suppl. material 3).

Information related to morphological characters and ploidy levels were collected from several sources and presented in Suppl. material 3. We selected six characters based on previous studies: number of anal and dorsal fin rays, number of barbels, presence or absence of barbels, length, ploidy levels, and type of lips (Howes 1991, Skelton 2001; Zheng et al. 2010; Yang et al. 2015). Character states were tabulated and mapped using Mesquite 3.04 (Maddison and Maddison 2015) onto the parsimonious molecular phylogenetic tree.

Results

The length of the aligned COI matrix was 652 bp with the following base composition: A: 25.9 %, C: 26.8 %, G: 18.2 % and T: 29.1 %. The interspecific genetic distances (K2P) ranged from 0 to 0.30 (median = 0.15) and are larger than the intraspecific genetic distances (range: 0 – 0.02; median = 0.001; $p < 0.001$; Figure 1). This is indicative of a barcode gap in the COI dataset of the studied Cyprinidae. The existence of a barcode gap is further confirmed when we compared the lowest interspecific versus the furthest intraspecific distance (Figure 2). We found the optimised distance d = 0.015 suitable for species discrimination in the studied African Cyprinidae (Figure 3). Based on this threshold, the performance of COI varies with the method used (Table 1). The near neighbour method shows a discriminatory power of 92.1 %. The other two methods provide a lower performance of 88.2 % for the best close match (278 specimens out of 315) and 79.4 % with the BOLD method.

In addition, the result presented in Figure 4 shows that most nodes form robust monophyletic clades (red-coded nodes in Figure 4). The level of monophyly is further confirmed on Figure 5 which clearly indicates two distinct subfamilies (Cyprininae and Danioninae) and five tribes in the subfamily Cyprininae (Figure 5). The mapping of morphological characters and ploidy level on the phylogeny indicates that some characters are clearly clustered [e.g., number of anal soft rays and presence/absence of barbels for the tribe Smiliogastrini, the fish length (21–40 cm) for the tribe Labeonini and the tetraploidy for Barbinini; fig. 5].

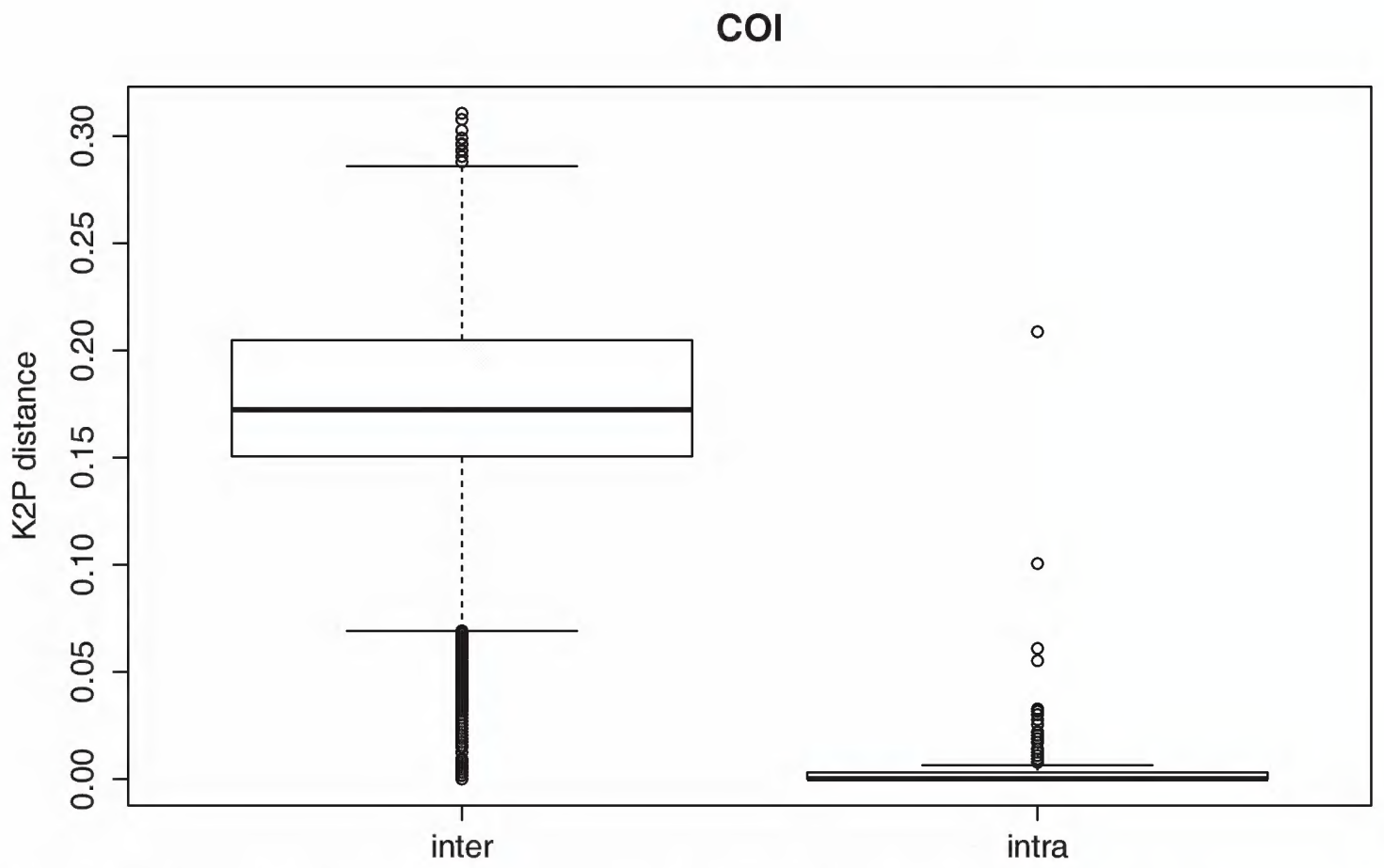

Figure 1. Evaluation of barcode gap in the dataset. Boxplot of the interspecific (inter) and intraspecific genetic (intra) distances, indicating the existence of a barcode gap, i.e., interspecific distance is larger than intraspecific distance. The median is indicated by the horizontal line and the range as the vertical dashed lines and outliers by bold vertical lines.

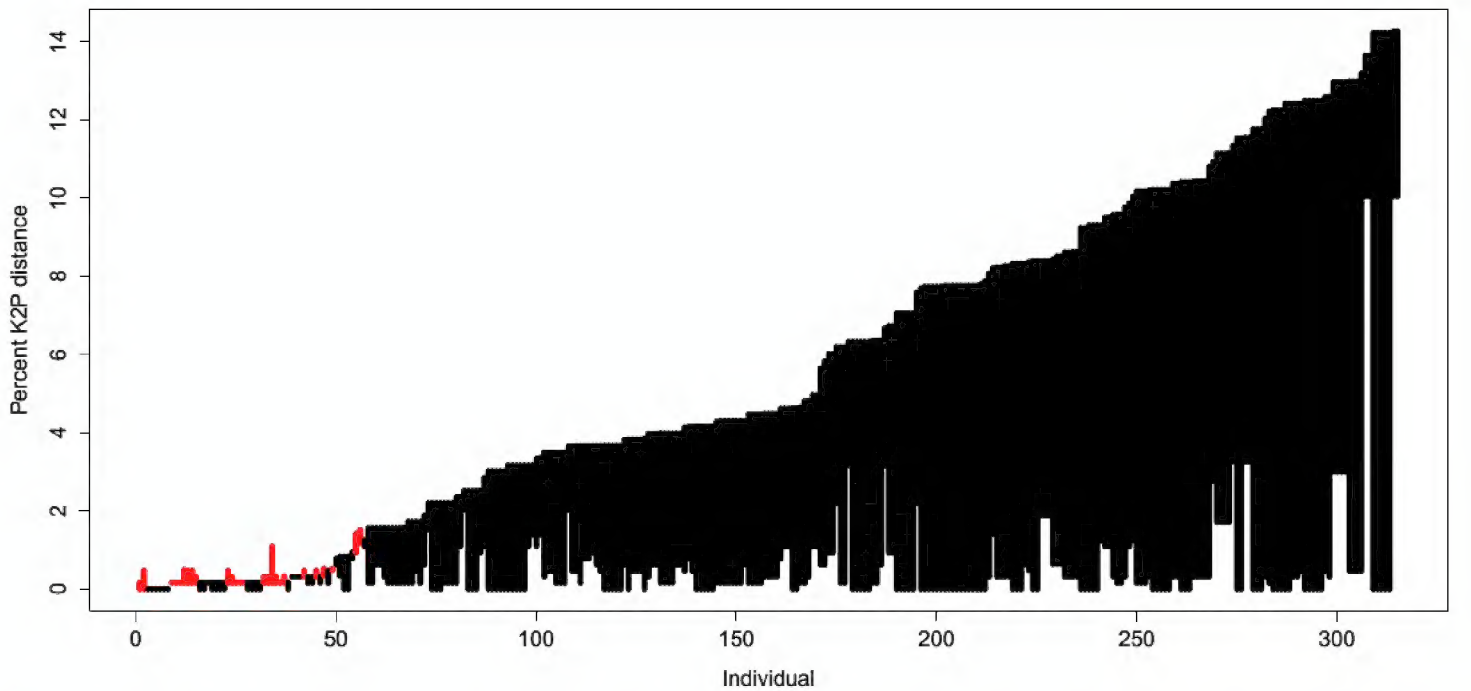

Figure 2. Evaluation of barcode gap in the dataset. Line plot of the barcode gap for the 315 Cyprinidae individuals. The black lines indicate where the smallest interspecific distance is longer than the longest intraspecific distance (bottom of line value), thus showing the existence of a barcode gap. The red lines show where this pattern is reversed.

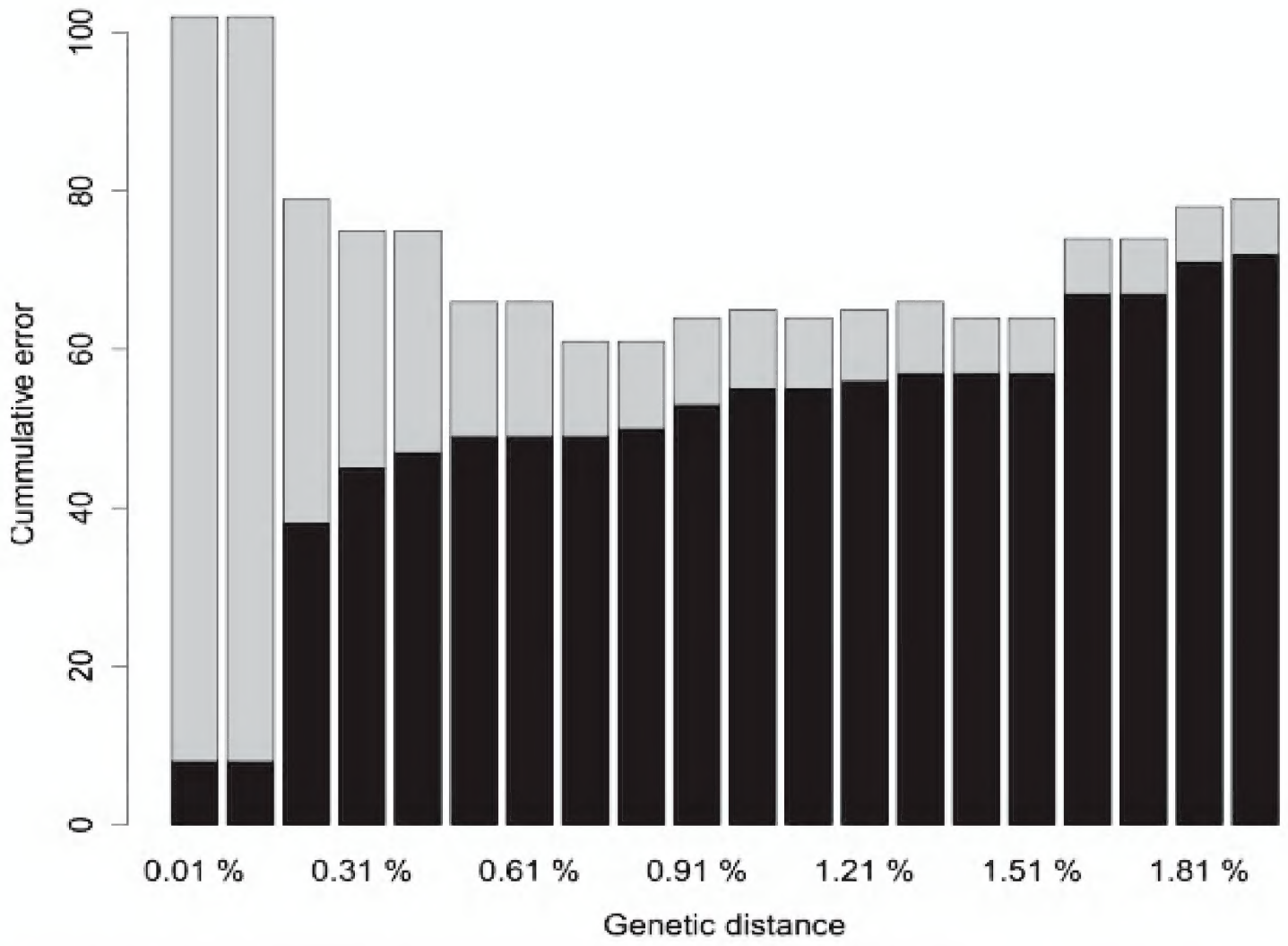

Figure 3. Determination of the threshold genetic distance for species identification. False positive (grey) and false negative (black) identification error rates summed across a range of distance thresholds from 0.01 to 1.9 %. The cumulative error plot indicates the transition between intraspecific and interspecific distances, the genetic distance corresponding to the least cumulative error (1.51 %) showing the appropriate threshold value for the dataset.

Discussion

Although COI is a universally accepted DNA barcode for animal groups (Hebert et al. 2003), its efficacy has also been questioned for some clades (Vences et al. 2005a; Chen et al. 2012; Murphy et al. 2013), and this prompts the need to assess its reliability for any particular group of interest (Collins et al. 2012).

The results presented in this work confirm that COI can be reliably used from a barcode perspective to distinguish between specimens of the African Cyprinidae in a dataset of 315 specimens representing 14 out of the 24 (58 %) recognised genera in Africa. For example, a significant barcode gap was found irrespective of the methods used, and this has also been reported for Cyprinidae of other geographic regions (e.g., Batishchevaa et al. 2011). Our results (79.4 %–92.1 %) from the distance-based method showed a pattern similar to the 90 % to 99 % discriminatory power reported for ornamental cyprinid fish species also mostly from Cyprininae and Danioninae and a catostomid (Collins et al. 2012). Irrespective of some drawbacks associated with the use of DNA barcoding and highlighted by some authors for some taxonomic groups (Vences et al. 2005a, b; Nielsen and Matz 2006; Valentini et al. 2008; Laskar et al.

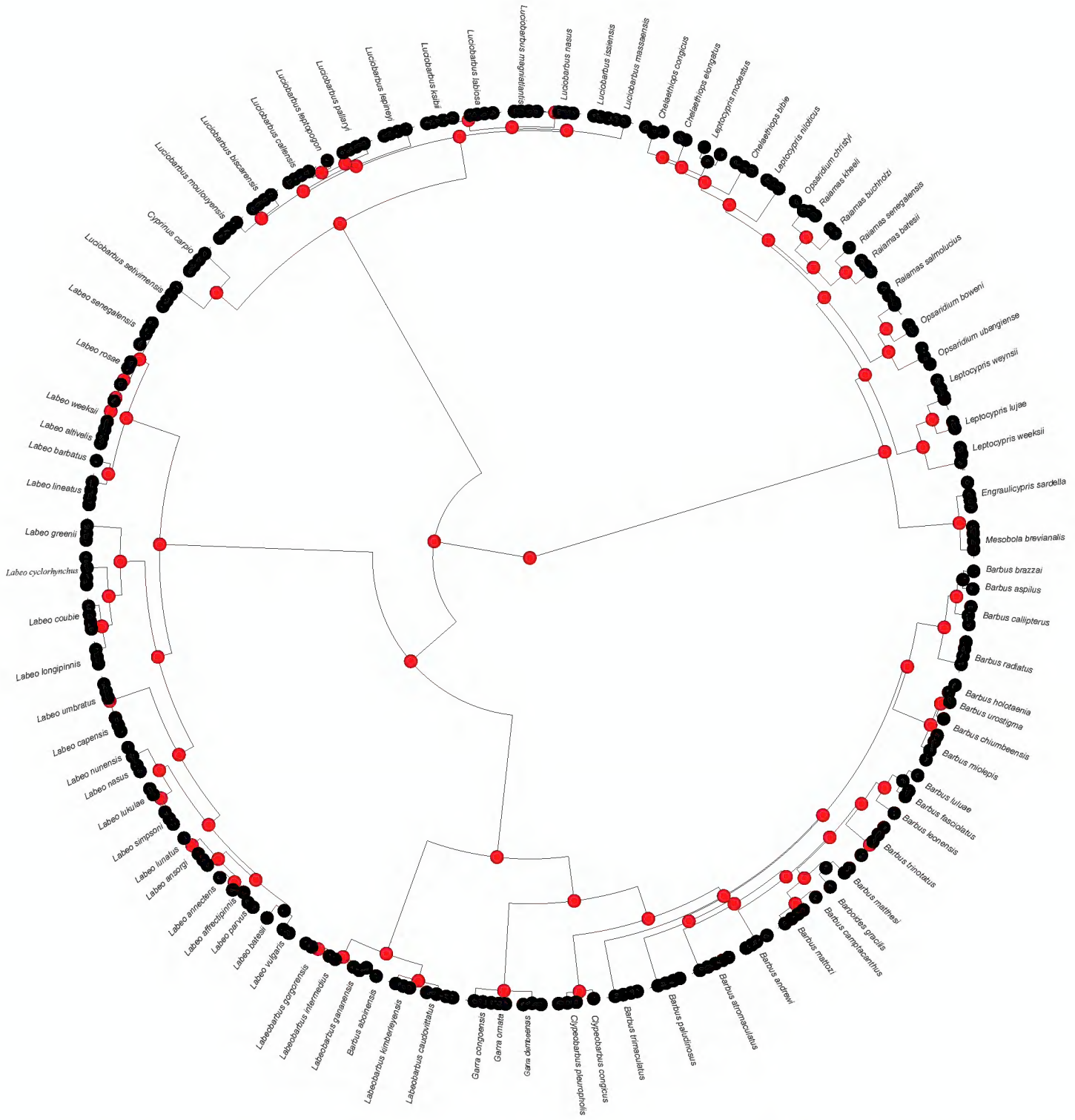

Figure 4. Neighbour-joining tree analysis using Rosenberg's (2007) test. Nodes in red are strongly supported nodes, indicating species monophyly.

2013) as well as the recent development of new generation sequencing techniques (e.g., Taylor and Harris 2012), the marker COI still remains useful for identification purposes (Batishchevaa et al. 2011; Collins et al. 2012; Van der Bank et al. 2013).

For example, the high level of COI discrimination is further supported by the test of species monophyly, a test that resulted in strongly supported clades based on Rosenberg (2007)'s probability of reciprocal monophyly on the NJ tree (see also Collins et al. 2012). Even our Maximum Parsimony tree provides additional support to the COI's power of discriminating between clades of the African Cyprinidae. Specifically, our phylogenetic analysis retrieved 14 monophyletic genera clearly grouped into two subfamilies (Cyprininae and Danioninae). Within the Cyprininae, five tribes are distinctly recovered: Barbini, Cyprinini, Labeonini, Smiliogastrini, and Torini as in

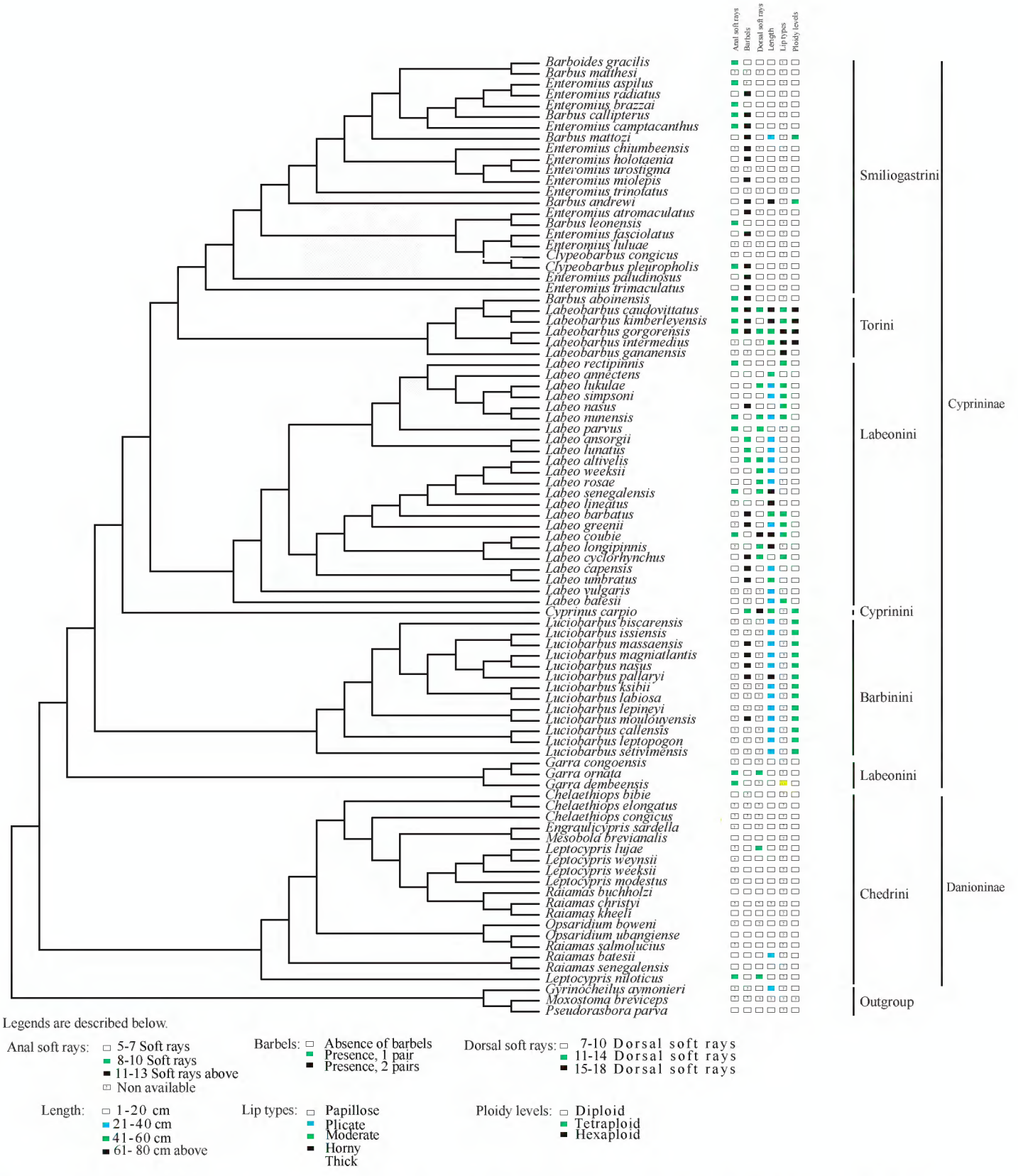

Figure 5. Parsimonious phylogenetic tree showing some plotted morphological characters, such as anal soft rays, barbels, dorsal soft rays, length, lip types, and ploidy status of 86 African Cyprinidae species.

Yang et al. (2015). The subfamily Danioninae was represented in our material by the tribe Chedrini which is well supported and includes *Chelaethiops*, *Engraulicyprus*, *Leptocypris*, *Mesobola*, *Opsaridium*, and *Raiamas* (see also Tang et al. 2010).

This evidence of monophyly accords with the morphology-based taxa delimitation as we found that some morphological characters and ploidy levels clustered within some clades along the phylogeny. Such characters that clustered within clades include, for example, the number of anal soft rays and presence/absence of barbels for the tribe Smiliogastrini, the fish length (21–40 cm) for the tribe Labeonini and the tetraploidy for Barbinini. Such clustering on the COI-phylogeny is evidence not only for COI as DNA barcoding of some

African Cyprinidae 11 a good barcode for the family Cyprinidae but also that COI-phylogeny can be used for a comparative phylogenetic analysis. Only the tribe Labeonini sensu Rainboth 1991 (Yang and Mayden 2010) was retrieved non-monophyletic in our dataset.

Overall, the existence of DNA barcode gap and a high discriminatory power, as well as the high level of monophyly give support to the use of COI as a reliable DNA barcode for African Cyprininae and Danioninae. Several studies have examined the phylogeny of this family at subfamily and genus levels using both mitochondrial and nuclear genes (Simons et al. 2003; Stiassny and Getahun 2007; Swartz et al. 2009; Tsigenopoulos et al. 2010; Yang et al. 2010; Zheng et al. 2012). Our study provides additional evidence for the effectiveness of DNA barcode data as a complementary tool to morphology-based identification of some African Cyprinidae, and our findings indicate that a COI-based phylogenetic tree for the African Cyprinidae can be used in comparative phylogenetic analyses and important applied problems (e.g., conservation) for this group of fish.

Acknowledgments

Kowiyou Yessoufou acknowledges the South Africa's National Research Foundation (NRF) for funding (Grant No: 103944) as well as La Société Botanique de France. He also acknowledges a start-up grant from the University of Johannesburg (URC Grant: 073450). We also thank the editor and anonymous reviewers for their comments.

References

Agnèse JF, Berrebi P, Lévêque C, Guégan JF (1990) Two lineages, diploid and tetraploid, demonstrated in African species of *Barbus* (Osteichtyes, Cyprinidae). Aquatic Living Resources 3: 305–311. https://doi.org/10.1051/alr:1990031

Armbruster JW, Stout CC, Hayes MM (2016) An empirical test for convergence using African barbs (Cypriniformes: Cyprinidae). Evolutionary Ecology 30: 435–450. http://dx.doi.org/10.1007/s10682-015-9811-6

Batishchevaa NM, Kartavtsev YP, Bogutskaya NG (2011) Phylogenetic analysis of Altai Osmans of the Genus *Oreoleuciscus* (Pisces, Cyprinidae, Leuciscinae), based on the analysis of the Cytochrome Oxidase 1 Gene (*COI*) Sequence. Russian Journal of Genetics 47: 1188–1197. http://dx.doi.org/10.1134/S1022795411100036

Berrebi P, Kottelat M, Skelton P, Rab P (1996) Systematics of *Barbus*: State of the art and heuristic comments. Folia Zoologica 45: 5–12.

Beshera KA, Harris PM, Mayden RL (2016) Novel evolutionary lineages in *Labeobarbus* (Cypriniformes; Cyprinidae) based on phylogenetic analyses of mtDNA sequences. Zootaxa 4093: 363–381. http://dx.doi.org/10.11646/zootaxa.4093.3.4

Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. Molecular Phylogenetic and Evolution 9: 100–108. https://doi.org/10.1006/mpev.1997.0441

Brown SDJ, Collins RA, Boyer S, Lefort MC, Malumbres-Olarte JA, Vink CJ, Cruickshank RH (2012) Spider: An R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Molecular Ecology Resources 12: 562–565. http://cran.r-project.org, http://spider.r-forge.rproject.org, https://doi.org/10.1111/j.1755-0998.2011.03108.x

Cavender TM, Coburn MM (1992) Phylogenetic relationships of North American Cyprinidae. In: Mayden RL (Ed.) Systematics, Historical ecology and North American Freshwater Fishes. Stanford University Press, Stanford, CA, 293–327.

Chen Z, Qi W, Yang J (2009) A morphologically variant natural population of cyprinids without dorsal fin possibly derived from Mystacoleucus marginatus (Osteichthyes: Teleostes) from the Lancang Jiang River, Yunnan, China. Zoologischer Anzeiger – A Journal of Comparative Zoology 248: 93–100. https://doi.org/10.1016/j.jcz.2009.02.002

Chen R, Jiang LY, Qiao GX (2012) The effectiveness of three regions in mitochondrial genome for aphid DNA barcoding: a case in Lachininae. PloS ONE 7: e46190. http://dx.doi.org/10.1371/journal.pone.0046190

Chen W, Ma X, Shen Y, Mao Y, He S (2015) The fish diversity in the upper reaches of the Salween River, Nujiang River, revealed by DNA barcoding. Scientific Reports 5: 17437. http://dx.doi.org/10.1038/srep17437

Collins RA, Armstrong KF, Meier R, Yi Y, Brown SDJ, Cruickshank RH, Keeling S, Johnston C (2012) Barcoding and border biosecurity: Identifying Cyprinid Fishes in the Aquarium Trade. PLoS ONE 7: e28381. http://dx.doi.org/10.1371/journal.pone.0028381

Decru E, Moelants T, DeGelas K, Vreven E, Verheven E, Snoeks J (2016) Taxonomic challenges in freshwater fishes: a mismatch between morphology and DNA barcoding in fish of the north-eastern part of the Congo Basin. Molecular Ecology Resources 16: 342–352. http://dx.doi.org/10.1111/1755-0998.12445

De Graaf M, Megens H, Samallo J, Sibbing FA (2007) Evolutionary origin of Lake Tana's (Ethiopia) small *Barbus* species: indications of rapid ecological divergence and speciation. Animal Biology 57: 39–48. http://dx.doi.org/10.1111/1755-0998.12445

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32: 1792–1797. http://dx.doi.org/10.1093/nar/gkh340

Eschmeyer WN, Fong JD (2015) Species by family/subfamily. http://research.calademy.org/research/ ichthyology/catalog/SpeciesByFamily.asp [Accessed 21 Feb 2016]

Froese R, Pauly D (2017) FishBase. http://www.fishbase.org [Accessed 4 March 2017]

Gilles A, Lecointre G, Miquelis A, Loerstcher M, Chappaz R, Brun G (2001) Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. Molecular Phylogenetic and Evolution 19: 22–33. https://doi.org/10.1006/mpev.2000.0916

Güegan JF, Rab P, Machordom A, Doadrio I (1995) New evidence of hexaploidy in 'large' African Barbus with some considerations on origin of hexaploidy. Journal of Fish Biology 47: 192–198. https://doi.org/10.1111/j.1095-8649.1995.tb01888.x

Hajibabaei M, De Waard JR, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie PM, Hebert PDN (2005) Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of the Royal Society B 360: 1959–1967. http://dx.doi.org/10.1098/rstb.2005.1727

He S, Mayden RL, Wang X, Wang W, Tang KL, Chen WJ (2008) Molecular phylogenetics of the family Cyprinidae (Actinopterygii: Cypriniformes) as evidenced by sequence variation in the

first intron of S7 ribosomal protein-coding gene: Further evidence from a nuclear gene of the systematic chaos in the family. Molecular Phylogenetics and Evolution 46: 818–829. http://dx.doi.org/10.1016/j.ympev.2007.06.001

Hebert PDN, Cywinska A, Ball SL, De Waard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society B 270: 313–321. http://dx.doi.org/10.1098/rspb.2002.2218

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLoS Biology 2: e312. https://doi.org/10.1371/journal.pbio.0020312

Howes GJ (1991) Systematics and biogeography: an overview. In: Winfield IJ (Ed.) Cyprinid Fishes – Systematics, Biology and Exploitation. Chapman & Hall, London, 1–33. https://doi.org/10.1007/978-94-011-3092-9_1

Imoto JM, Saitoh K, Sasaki T, Yonezawa T, Adachi J, Kartavtsev YP, Miya M, Nishida M, Hanzawa N (2013) Phylogeny and biogeography of highly diverged freshwater fish species (Leuciscinae, Cyprinidae, Teleostei) inferred from mitochondrial genome analysis. Gene 514: 112–124. http://dx.doi.org/10.1016/j.gene.2012.10.019

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120. http://dx.doi.org/10.1007/BFo131581

Kochzius M, Seidel C, Antoniou A, Botla SK, Campo D, Botla SD, Cariani A, Vazque EG, Hauschild J, Herve C, Hjörleifsdottir S, Hreggvidsson G, Kappel K, Landi M, Magoulas A, Marteinsson V, Nölte M, Planes S, Tinti F, Turan C, Venugopal MN, Weber H, Blohm D (2010) Identifying fishes through DNA barcodes and microarrays. PLoS ONE 5: e12620. http://dx.doi.org/10.1371/journal.pone.0012620

Laskar BA, Bhattacharjee MJ, Dhar B, Mahadani P, Kundu S, Ghosh SK (2013) The Species Dilemma of Northeast Indian Mahseer (Actinopterygii: Cyprinidae): DNA Barcoding in Clarifying the Riddle. PloS ONE 8: e53704. http://dx.doi.org/10.1371/journal.pone.0053704

Liao TY, Kullander SO, Fang F (2011) Phylogenetic position of rasborin cyprinids and monophyly of major lineages among the Danioninae, based on morphological characters (Cypriniformes: Cyprinidae). Journal of Zoological Systematics and Evolutionary Research 49: 224–232. http://10.1111/j.1439-0469.2011.00621.x

Little DP (2011) DNA Barcode sequence identification incorporating taxonomic hierarchy and within taxon variability. PLoS ONE 6: e20552. http://dx.doi.org/10.1371/journal.pone.0020552

Machordom A, Doadrio I (2001) Evolutionary history and speciation modes in the cyprinid genus *Barbus*. Proceedings of the Royal Society London. Biological Science 268: 1297–1306. http://dx.doi.org/10.1098/rspb.2001.1654

Maddison WP, Maddison DR (2015) Mesquite: A modular system for evolutionary analysis. Version 3.04, build 725. http://mesquiteproject.org.

Meier R, Shiyang K, Vaidya G, Ng PK (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Systematic Biology 55: 715–728. http://dx.doi.org/10.1080/10635150802406343

Meier R, Zhang G, Ali F (2008) The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. Systematic Biology 57: 809–813. http://dx.doi.org/10.1371/journal.pbio.0030422

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biology 3: e422. http://dx.doi.org/10.1371/journal.pbio.0030422

Murphy RW, Crawford AJ, Bauer AM, Che J, Donnellan SC, Fritz U, Haddad CF, Nagy ZT, Poyarkov NA, Vences M, Wang WZ (2013) Cold Code: the global initiative to DNA barcode amphibians and non-avian reptiles. Molecular Ecology Resources 13: 161–167. http://dx.doi.org/10.1111/1755-0998.12050

Nelson JS (2006) Fishes of the World, Fourth edition. John Wiley, New York, 624 pp.

Nelson JS, Grande TC, Wilson MV (2016) Fishes of the World. John Wiley & Sons, 634 pp. https://doi.org/10.1002/9781119174844

Nielsen R, Matz M (2006) Statistical approaches for DNA barcoding. Systematic Biology 55: 162–169. http://dx.doi.org/10.1080/10635150500431239

Nigro LM, Angel MV, Blachowiak-Samolyk K, Hopcroft RR, Bucklin A (2016) Identification, discrimination, and discovery of species of marine planktonic Ostracods using DNA Barcodes. PLoS ONE 11: e0146327. http://dx.doi.org/10.1371/journal.pone.0146327

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, Kõljalg U (2006) Taxonomic reliability of DNA sequences in public sequence databases: A fungal perspective. PLoS ONE 1: e59. http://dx.doi.org/10.1371/journal.pone.0000059

Oellermann LK, Skelton PH (1990) Hexaploidy in yellowfish species (*Barbus*, Pisces, Cyprinidae) from southern Africa. Journal of Fish Biology 37: 105–115. https://doi.org/10.1111/j.1095-8649.1990.tb05932.x

Pereira LHG, Maia GMG, Hanner R, Foresti F, Oliveira C (2011) DNA barcodes discriminate freshwater fishes from the Paraiba do Sul River Basin, São Paulo, Brazil. Mitochondrial DNA 22: 71–79. http://dx.doi.org/10.3109/19401736.2010.532213

Ráb P, Collares-Pereira MJ (1995) Chromosomes of European cyprinid fishes (Cyprinidae, Cypriniformes): a review. Folia Zoology 44: 193–214.

Rainboth WJ, Vidthayanon C, Mai DY (2012) Fishes of the Greater Mekong Ecosystem with Species List and Photographic Atlas, vol. 201. Miscellaneous Publications, Museum of Zoology, University of Michigan, Berlin (i–vi, 1–173).

Ren Q, Mayden RL (2016) Molecular phylogeny and biogeography of African diploid barbs, '*Barbus*', and allies in Africa and Asia (Teleostei: Cypriniformes). Zoologica Scripta 45: 642–649. http://dx.doi.org/10.1111/zsc.12177

Rosenberg NA (2007) Statistical tests for taxonomic distinctiveness from observations of monophyly. Evolution 61: 317–323. http://dx.doi.org/10.1111/j.1558-5646.2007.00023.x

Schmidt RC, Bart HL (2015) Nomenclatural changes should not be based on equivocally supported phylogenies: Reply to Yang et al. 2015. Molecular Phylogenetics and Evolution 90: 193–194. http://dx.doi.org/10.1016/j.ympev.2015.05.025

Sethusa MT, Millar IM, Yessoufou K, Jacobs A, Van der Bank M, Van der Bank H (2014) DNA barcode efficacy for the identification of economically important scale insects (Hemiptera: Coccoidea) in South Africa. African Entomology 22: 257–266. http://dx.doi.org/10.4001/003.022.0218

Simons AM, Berendzen PB, Mayden RL (2003) Molecular systematics of North American Phoxinin genera (Actinopterygii: Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. Zoological Journal of the Linnean Society 139: 63–80. http://dx.doi.org/10.1046/j.1096-3642.2003.00076

Skelton PH (2016) Name changes and additions to the southern African freshwater fish fauna. African Journal of Aquatic Science 41: 345–351. http://dx.doi.org/10.2989/16085914.2016.1186004

Skelton PH (2001) A complete guide for fresh water fishes of southern Africa. Struik Publishers, Cape Town, 395 pp.

Skelton PH, Tweddle D, Jackson PBN (1991) Cyprinids in Africa. In: Winfield IJ, Nelson JS (Eds) Cyprinid Fishes–Systematics, Biology and Exploitation. Chapman & Hall, London, 211–239. https://doi.org/10.1007/978-94-011-3092-9_7

Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov EV (2014) Is DNA barcoding actually cheaper and faster than traditional morphological methods: results from a survey of freshwater bioassessment efforts in the United States? PLoS ONE 9: e95525. http://dx.doi.org/10.1371/journal.pone.0095525

Steinke D, Hanner R (2011) The FISH-BOL collaborators' protocol. Mitochondrial DNA 22: 10–14. https://doi.org/10.3109/19401736.2010.536538

Stiassny MLJ, Getahun A (2007) An overview of Labeonini relationships and the phylogenetic placement of the Afro-Asian genus *Garra* Hamilton, 1922 (Teleostei: Cyprinidae), with the description of five new species of *Garra* from Ethiopia and a key to all African species. Zoological Journal of the Linnean Society. 150: 41–83. http://dx.doi.org/10.1111/j.1096-3642.2007.00281

Stiassny MLJ, Schelly RC, Schliewen UK (2006) A new species of Raiamas (Teleostei: Cyprinidae) from the Lower Congo River, with a phylogenetic assessment of the generic limits of the predatory cyprinid genera Opsaridium, Raiamas and Leptocypris. Copeia 2006: 370–377. https://doi.org/10.1643/0045-8511(2006)2006[370:ANSORT]2.0.CO;2

Swartz ER, Skelton PH, Bloomer P (2009) Phylogeny and biogeography of the genus *Pseudobarbus* (Cyprinidae): Shedding light on the drainage history of rivers associated with the Cape Floristic Region. Molecular Phylogenetics and Evolution 51: 75–84. http://dx.doi.org/10.1016/j.ympev.2008.10.017

Swofford DL (2002) PAUP: Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts. https://doi.org/10.1111/j.0014-3820.2002.tb00191.x

Tang KL, Agnew MK, Hirt MV, Sado T, Schneider LM, Freyhof J, Sulaiman Z, Swart E, Vidthavanon C, Miya M, Saitoh K, Soimios AM, Wood RM, Mayden RL (2010) Systematics of the subfamily Danioninae (Teleostei: Cypriniformes: Cyprinidae). Molecular Phylogenetics and Evolution 57: 189–214. http://dx.doi.org/10.1016/j.ympev.2010.05.021

Taylor HR, Harris WE (2012) An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. Molecular Ecology Resources 12: 377–388. http://dx.doi.org/10.1111/j.1755-0998.2012.03119.x

Thai BT, Phan PD, Austin CM (2007) Phylogenetic evaluation of subfamily classification of the Cyprinidae focusing on Vietnamese species. Aquatic Living Resources 20: 143–153. http://dx.doi.org/10.1051/alr:2007025

Tsigenopoulos CS, Kasapidis P, Berrebi P (2010) Phylogenetic relationships of hexaploid large-sized barbs (genus *Labeobarbus*, Cyprinidae) based on mtDNA data. Molecular Phylogenetics and Evolution 56: 851–856. http://dx.doi.org/10.1016/j.ympev.2010.02.006

Tsigenopoulos CS, Rab P, Naran D, Berrebi P (2002) Multiple origins of polyploidy in the phylogeny of southern African barbs (Cyprinidae) as inferred from mtDNA markers. Heredity 88: 466–473. https://doi.org/10.1038/sj.hdy.6800080

Valentini A, Pompanon F, Taberlet P (2008) DNA barcoding for ecologists. Trends in Ecology and Evolution 24: 110–117. http://dx.doi.org/10.1016/j.tree.2008.09.011

Van der Bank FH, Herbert D, Greenfield R, Yessoufou K (2013) Revisiting species delimitation within the genus *Oxystele* using DNA barcoding approach. ZooKeys 365: 337–354. http://dx.doi.org/10.3897/zookeys.365.5356

Vences M, Thomas M, Van der Meijden A, Chiari Y, Vieites DR (2005a) Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. Frontiers in Zoology 2: 1–12. http://dx.doi.org/10.1186/1742-9994-2-5

Vences M, Thomas M, Bonett RM, Vieites DR (2005b) Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philosophical Transactions of the Royal Society London B Biological Science 360: 1859–1868. http://dx.dx.doi.org/10.098/rstb.2005.1717

Vreven EJWMN, Musschoot T, Snoeks J, Schliewen UK (2016) The African hexaploid Torini (Cypriniformes: Cyprinidae): review of a tumultuous history. Zoological Journal of the Linnean Society 177: 231–305. http://dx.doi.org/10.1111/zoj.12366

Ward RD, Zemlak TS, Innes BH, Last PD, Hebert PDN (2005) DNA Barcoding Australia fish species. Philosophical Transactions of the Royal Society B 360: 1847–1857. http://dx.org/10.1098/rstb.2005.1716

Yang L, Mayden RL (2010) Phylogenetic relationships, subdivision, and biogeography of the cyprinid tribe Labeonini (sensu)(Teleostei: Cypriniformes), with comments on the implications of lips and associated structures in the labeonin classification. Molecular Phylogenetics and Evolution 54: 254–265. http://dx.doi.org/10.1016/j.ympev.2009.09.027

Yang L, Mayden RL, Sado T, He SP, Saitoh K, Miya M (2010) Molecular phylogeny of the fishes traditionally referred to Cyprinini sensu stricto (Teleostei: Cypriniformes). Zoological Scripta 39: 527–550. http://dx.doi.org/10.1111/j.1463-6409.2010.00443.x

Yang L, Sado T, Hirt MV, Pasco-Viel E, Arunachalam M, Li J, Wang X, Freyhof J, Saitoh K, Simons AM, Miya M, He S, Mayden RL (2015) Phlogeny and polyploidy: Resolving the classification of cyprinine fishes (Teleostei: Cypriformes). Molecular Phlogenetics and Evolution 85: 97–116. http://dx.doi.org/10.1016/j.ympev.2015.01.014

Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European Cyprinids. Journal of Molecular Evolution 49: 227–237. https://doi.org/10.1007/PL00006545

Zheng LP, Yang JX, Chen XY, Wang WY (2010) Phylogenetic relationships of the Chinese Labeoninae (Teleostei, Cypriniformes) derived from two nuclear and three mitochondrial genes. Zoologica Scripta 39: 559–571. http://dx.doi.org/10.1111/j.1463-6409.2010.00441.x

Zheng L, Yang H, Chen X (2012) Phylogeny of the Labeoninae (Teleostei, Cypriniformes) based on nuclear DNA sequences and implications on character evolution and biogeography. Current Zoology 58: 8379–850. http://dx.doi.org/10.1093/czoolo/58.6.837

Zhou W (1989) Phylogeny of the subfamily Cyprininae (Pisces: Cyprinidae). Acta Zootaxonomica Sinica 14: 247–256. [In Chinese with English abstract]

Supplementary material 1

Table S1

Authors: Adeoba MI, Kabongo R, Van der Bank H, Yessoufou K

Data type: occurence

Explanation note: Taxonomy of African Cyprinidae and geographic origins.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.746.13502.suppl1

Supplementary material 2

Table S2

Authors: Adeoba MI, Kabongo R, Van der Bank H, Yessoufou K

Data type: occurence

Explanation note: List of 315 Cyprinidae specimens analysed in this study. Full names, voucher information, and geographic origins are presented. Accession numbers for ACDB (African Centre for DNA Barcoding at the University of Johannesburg), GenBank, and BOLD are also included.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.746.13502.suppl2

Supplementary material 3

Table S3

Authors: Adeoba MI, Kabongo R, Van der Bank H, Yessoufou K

Data type: measurement

Explanation note: Matrix of the morphological characters used for character mapping along a phylogeny. Details of the characters (1–6) and sources are provided at the bottom of the table..

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.746.13502.suppl3